A SIMPLE SHARP WEIGHTED ESTIMATE OF THE DYADIC SHIFTS ON METRIC SPACES WITH GEOMETRIC DOUBLING

FEDOR NAZAROV AND ALEXANDER VOLBERG

ABSTRACT. We give a short and simple polynomial estimate of the norm of weighted dyadic shift on metric space with geometric doubling, which is linear in the norm of the weight. Combined with the existence of special probability space of dyadic lattices built in [27], and with decomposition of Calderón–Zygmund operators to dyadic shifts from [4] (and later [8]), we will be able to have a linear (in the norm of weight) estimate of an arbitrary Calderón–Zygmund operator on a metric space with geometric doubling. This will be published separately.

1. INTRODUCTION

Recall that in [24] it was proved that

Theorem 1.1. *If T is an arbitrary operator with a Calderón–Zygmund kernel, then*

$$\|T\|_{L^2(wd\mu)\to L^{2,\infty}(w)}+\|T'\|_{L^2(w^{-1})\to L^{2,\infty}(w^{-1})} \le 2\|T\|_{L^2(wd\mu)\to L^2(wd\mu)}$$
$$\le C\left([w]_{A_2} + \|T\|_{L^2(wd\mu)\to L^{2,\infty}(w)} + \|T'\|_{L^2(w^{-1})\to L^{2,\infty}(w^{-1})}\right).$$

By T' we denote the adjoint operator. Here of course only the right inequality is interesting. And it is unexpected too. The weak and strong norm of any operator with a Calderón–Zygmund kernel turned out to be equivalent up to additive term $[w]_{A_2}$. In its turn, Theorem 1.1 was obtained in [24] as a corollary of a weighted $T1$ theorem–the Main Theorem of [24]. Again in its turn the Main Theorem (=*weighted $T1$ theorem*) in [24] is a consequence of a rather difficult two-weight $T1$ theorem of [16]. From Theorem 1.1 we obtained in [24] the following result which holds for any Calderón–Zygmund operator.

Theorem 1.2. $\|T\|_{L^2(wd\mu)\to L^2(wd\mu)} \le C \cdot [w]_{A_2} \log(1 + [w]_{A_2})$.

By A_2 conjecture people understand the strengthening of this claim, where the logarithmic term is deleted, in other words, a linear (in weight's norm) estimate of arbitrary weighted Calderón–Zygmund operator. In [6] the A_2 conjecture was proved for Calderón–Zygmund operators having more than $2d$ smoothness in $\mathbb{R}^d$.

A bit later a preprint [4] of Tuomas Hytönen has appeared, the A_2 conjecture is fully proved there. It is based on the Main Theorem (=*weighted $T1$ theorem*) in [24] of Pérez–Treil–Volberg. Both [24] and [4] are neither short nor easy.

Notice that the scheme of the proof in [4] goes like that: given a Calderón–Zygmund operator T and a $w \in A_2$ weight, one first uses the Main Theorem (=*weighted $T1$ theorem*) of [24], which says that to prove a linear estimate for $\|Tf\|_w$ it is enough to prove it uniformly only for special "characteristic functions of cubes" type functions f (see the true statement in [24]). Notice that all the cubes must be checked. The second (very beautiful) part of the proof is to decompose T into "a convex combinations" of dyadic shifts, the new idea is used here that grew out of random lattices approach

in non-homogeneous Harmonic Analysis theory of Nazarov–Treil–Volberg, see [13], [15], [16], [29].
Now it is enough to check the right estimate for each dyadic shift applied to each "characteristic
function of cube". A very annoying difficulty appears: the shift is with respect to a certain dyadic
lattice, but the cube in question is arbitrary and a priori does not belong to this lattice. This creates
serious technical difficulties, which can be (and were) avoided in [8].

The direct proof of A_2 conjecture (without going through [24]) was given in [8], and it was based
on two ingredients: 1) a formula for decomposing an arbitrary Calderón–Zygmund operators into
(generalized) dyadic shifts by the averaging trick, see [4], 2) on a polynomial in complexity and
linear in weight estimate of the norm of a dyadic shift.

The latter was quite complicated and was based on modification of the argument in Lacey–Petermichl-
Reguera [9] and on the use of [14] with its careful reexamination. The former–as we already mentioned–
was rooted in the works on non-homogeneous Harmonic Analysis, like e. g. [13]–[15], [16], [29], but
with a new twist, which appeared first in Hytönen's [4] and was simplified in Hytönen–Pérez–Treil–
Volberg's [8].

The averaging trick was a development of the bootstrapping argument used by Nazarov–Treil–
Volberg, where they exploited the fact that the bad part of a function can be made arbitrarily small.
Using the original Nazarov–Treil–Volberg averaging trick would add an extra factor depending on
$[w]_{A_2}$ to the estimate, so a new idea was necessary. The new observation in [4] was that as soon as the
probability of a "bad" cube is less than 1, it is possible to completely ignore the bad cubes (at least in
the situation where they cause troubles).

2. SHIFTS OF COMPLEXITY m,n

Let X be a compact geometrically doubling metric space. Let μ be a doubling measure on X,
which exists by [10]. Let $\mathscr{D}$ be a finite "dyadic" lattice on X. Namely, $\mathscr{D}$ consists of disjoint partition
$\mathscr{E}_N$ to pieces Q_N^j of size $\asymp \delta^N$, then there is a partition $\mathscr{E}_{N-1}$ to pieces Q_{N-1}^i of size δ^{N-1}, each Q_{N-1}^i
consists of boundedly many (at most $M_{N-1}(i) \leq M$) pieces of partition $\mathscr{E}_N$, et cetera... , we have
$\mathscr{E}_k, k = N, N-1, ..., 0$, and each Q_k^i is an almost ball: if $\ell(Q), Q = Q_k^i$, denote its diameter, then there
is a ball of radius $c\ell(Q), c > 0$, inside Q, and c does not depend on N, k, i.

The existence of such lattices with all constants depending only on geometric doubling of X was
proved by Christ [3].

We relate to $\mathscr{E}_k$ the projection operator $\mathbb{E}_k$ on $L^2(\mu)$: $\mathbb{E}_k f = \sum_j \langle f \rangle_{Q_k^j, \mu} \chi_{Q_k^j}$. We also consider the
martingale difference operator $\Delta_k f = E_k f - E_{k-1} f$. Notice that it can be written as

$$\Delta_k f = \sum_i \sum_{j=1}^{M_{k-1}(Q_{k-1}^i)} (f, h_{Q_{k-1}^i}^j)_\mu h_{Q_{k-1}^i}^j ,$$

where denoting $Q = Q_{k-1}^i$ we notice that $M_{k-1}(Q) \leq M$, and h_Q^j, are functions supported on Q, or-
thogonal to constants in $L^2(\mu)$, orthogonal to each other in $L^2(\mu)$, constant on each $Q_k^i \subset Q, Q_k^i \in \mathscr{E}_k$
(such Q_k^i are called the **sons** of Q_{k-1}^j), and having the following bound

$$(2.1) \qquad\qquad\qquad \|h_Q^j\|_{L^\infty(\mu)} \leq \frac{C}{\mu(Q)^{1/2}} .$$

Definition. We call such $h_{Q_k}^j$ Haar functions. If $L \in \mathscr{E}_m, J \in \mathscr{E}_{m+n}$ we say $g(J) = g(L) + n$.

We always use $\ell(S)$ to denote the diameter of a set S. Christ's cubes will be denoted Q, I, J, L, K,
may be with indices.

We call by $\mathbb{S}_{m,n}$ the operator given by the kernel

$$f \to \sum_{L \in \mathscr{D}} \int_L a_L(x,y) f(y) dy,$$

where

$$a_L(x,y) = \sum_{\substack{I \subset L, J \subset L \\ g(I)=g(L)+m, g(J)=g(L)+n}} c_{L,I,J} h_J^j(x) h_I^i(y),$$

where h_I^i, h_J^j are Haar functions normalized in $L^2(d\mu)$ and satisfying (2.1), and $|c_{L,I,J}| \le \frac{\sqrt{\mu(I)}\sqrt{\mu(J)}}{\mu(L)}$. Often we will skip superscripts i, j.

We are interested in sufficiently good estimate of

$$\|\mathbb{S}_{m,n}\|_w := \|\mathbb{S}_{m,n} : L^2(wd\mu) \to L^2(wd\mu)\|,$$

where $w \in A_2$. For such w we put $\sigma = w^{-1}$ and

$$[w]_{A_2} := \sup_I \langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I} < \infty,$$

and call it the norm of w (it is not a norm).

In recent paper [8] the following theorem was proved (another proof, using the Bellman function technique, was given recently in [12])

Theorem 2.1.

(2.2) $$\|\mathbb{S}_{m,n}\|_w \le C (m+n+1)^a [w]_{A_2}.$$

In [8] $a = 3$. Looks like here we have the same numerical value. But for its main application in [8]: the proof of A_2 conjecture, the value of a (if finite) is not important. The proof was hard and combinatorial, it was based on the ideas of [9], where such an estimate was proved with exponential dependence on $m+n$. We propose here a simple proof based on Bellman function technique. This technique was tried successfully for shifts of low complexity, first in [30], [20] then in [21], [22], and recently in preprint [26], which gives a simple unified proof of results in [30], [21], [22]. As the reader will see one needs a couple of new tricks to achieve this fulfillment. The proof below is a direct and simple readjustment of the proof in [12], where it has been carried out if $X = \mathbb{R}$, $\mu = dx$.

Remark. The reasoning below is in $\mathbb{R}$. But one can modify it without any efforts to any $\mathbb{R}^d$. Moreover, in [27] the probability space of Christ's type dyadic lattices is built on any compact metric space with the property of geometric doubling (every ball contains at most a fixed number of disjoint balls of half a radius), which allows to extend the sharp bound of Calderón–Zygmund operators into metric space setting by repeating the averaging trick that reduces everything to the case of dyadic shift on the metric space, and then using this preprint to give a polynomial in complexity and linear in weight estimate for any shift.

3. THE HEART OF THE MATTER: A REDUCTION TO BILINEAR EMBEDDING ESTIMATE

To prove Theorem 2.1 we need the following decomposition:

Lemma 3.1.

$$h_I^j = \alpha_I^j h_I^{w,j} + \beta_I^j \chi_I,$$

where
1) $|\alpha_I^j| \le \sqrt{\langle w \rangle_{\mu,I}}$,
2) $|\beta_I^j| \le \frac{(h_I^{w,j}, w)_\mu}{w(I)}$, where $w(I) := \int_I w d\mu$,

3) $\{h_I^{w,j}\}_I$ is supported on I, orthogonal to constants in $L^2(w\,d\mu)$,

4) $h_I^{w,j}$ assumes on each son $s(I)$ a constant values,

5) $\|h_I^{w,j}\|_{L^2(\mu)} = 1$.

Definition. Let
$$\Delta_I w := \sum_{\text{sons of } I} |\langle w \rangle_{\mu,s(I)} - \langle w \rangle_{\mu,I}|.$$

Remark. There are many Δ's in this paper, but the reader should notice that Δ_k is an operator sending functions from $L^2(\mu)$ to such functions with extra properties of being constants on each element of $\mathscr{E}_k$ and also being orthogonal in $L^2(\mu)$ to χ_L for any element L of $\mathscr{E}_{k-1}$. On the other hand, $\Delta_I w$ (as well as $\Delta_J \sigma$) is a non-negative number.

It is a easy to see (by a small linear algebra reasoning) that the doubling property of measure μ implies

(3.1) $$|(h_I^{w,j}, w)_\mu| \leq C \cdot (\Delta_I w) \sqrt{\mu(I)}.$$

Here C depends only on the doubling constant of measure μ. In other words, taking into account that $\Delta_k w = \sum_{I \in \mathscr{E}_{k-1}} \sum_{j=1}^{M_{k-1}(I)} (w, h_I^{w,j})_\mu h_I^{w,j}$ and given $I \in \mathscr{E}_{k-1}$, we can rewrite (3.1) as follows

(3.2) $$\chi_I |\Delta_k w| \leq C \cdot (\Delta_I w),$$

where C again depends only on the doubling constant of μ.

Therefore, the property 2) above can be rewritten as

2') $|\beta_I^j| \leq C \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{1}{\sqrt{\mu(I)}}$.

Fix $\phi \in L^2(w\,d\mu)$, $\psi \in L^2(\sigma)$. We need to prove

(3.3) $$|(\mathbb{S}_{m,n}\phi w, \psi \sigma)| \leq C(n+m+1)^a \|\phi\|_w \|\psi\|_\sigma.$$

We estimate $(\mathbb{S}_{m,n}\phi w, \psi \sigma)$ as (we skip superscripts j, i and write just h_J, h_I):

$$|\sum_L \sum_{I,J} c_{L,I,J}(\phi w, h_I)_\mu (\psi \sigma, h_J)_\mu| \leq$$

$$\sum_L \sum_{I,J} |c_{L,I,J}(\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} (\psi \sigma, h_J^\sigma)_\mu| \sqrt{\langle \sigma \rangle_{\mu,J}}| +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \phi w \rangle_{\mu,I} \frac{\Delta_I w}{\langle w \rangle_{\mu,I}} (\psi \sigma, h_J^\sigma)_\mu \sqrt{\langle \sigma \rangle_{\mu,J}} \sqrt{\mu(I)}| +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \psi \sigma \rangle_{\mu,J} \frac{\Delta_J \sigma}{\langle \sigma \rangle_{\mu,J}} (\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} \sqrt{\mu(J)}| +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \phi w \rangle_{\mu,I} \langle \psi \sigma \rangle_{\mu,J} \frac{\Delta_I w}{\langle w \rangle_{\mu,I}} \frac{\Delta_J \sigma}{\langle \sigma \rangle_{\mu,J}} \sqrt{\mu(I)} \sqrt{\mu(J)}| =: I + II + III + IV.$$

We can notice that because we have $|c_{L,I,J}| \leq \frac{\sqrt{\mu(I)}\sqrt{\mu(J)}}{\mu(L)}$, each sum inside L can be estimated by a perfect product of S and R terms, where

$$R_L(\phi w) := \sum_{I \subset L\dots} \langle \phi w \rangle_{\mu,I} \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}}$$

$$S_L(\phi w) := \sum_{I \subset L\dots} (\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} \frac{\sqrt{\mu(I)}}{\sqrt{\mu(L)}}$$

and the corresponding terms for $\psi\sigma$. So we have

$$I \leq \sum_L S_L(\phi w)S_L(\psi\sigma), II \leq \sum_L S_L(\phi w)R_L(\psi\sigma),$$

$$III \leq \sum_L R_L(\phi w)S_L(\psi\sigma), \ IV \leq \sum_L R_L(\phi w)R_L(\psi\sigma).$$

Now

(3.4) $$S_L(\phi w) \leq \sqrt{\sum_{I \subset L...} |(\phi w, h_I^w)_\mu|^2}\sqrt{\langle w\rangle_{\mu,L}}, \ S_L(\psi\sigma) \leq \sqrt{\sum_{J \subset L...} |(\psi\sigma, h_J^\sigma)|^2}\sqrt{\langle\sigma\rangle_{\mu,L}}$$

Therefore,

(3.5) $$I \leq C[w]_{A_2}^{1/2}\|\phi\|_w\|\psi\|_\sigma.$$

Terms II, III are symmetric, so consider III. Using Bellman function $(xy)^\alpha$ one can prove now

Lemma 3.2. *The sequence*

$$\mu_I := \langle w\rangle_{\mu,I}^\alpha\langle\sigma\rangle_{\mu,I}^\alpha\left(\frac{|\Delta_I w|^2}{\langle w\rangle_{\mu,I}^2} + \frac{|\Delta_I \sigma|^2}{\langle\sigma\rangle_{\mu,I}^2}\right)\mu(I)$$

form a Carleson measure with Carleson constant at most $c_\alpha Q^\alpha$, where $Q := [w]_{A_2}$ for any $\alpha \in (0,1/2)$. Here c_α depends only on α and the doubling constant of μ (and is independent of the doubling constants of $w\,d\mu$, $\sigma\,d\mu$).

Proof. We need a very simple

Sublemma. Let $Q > 1, 0 < \alpha < \frac{1}{2}$. In domain $\Omega_Q := \{(x,y) : X > o, y > 0, 1 < xy \leq Q$ function $B_Q(x,y) := x^\alpha y^\alpha\}$ satisfies the following estimate of its Hessian matrix (of its second differential form, actually)

$$-d^2 B_Q(x,y) \geq \alpha(1-2\alpha)x^\alpha y^\alpha\left(\frac{(dx)^2}{x^2} + \frac{(dy)^2}{y^2}\right).$$

The form $-d^2 B_Q(x,y) \geq 0$ everywhere in $x > 0, y > 0$. Also obviously $0 \leq B_Q(x,y) \leq Q^\alpha$ in Ω_Q.

Proof. Direct calculation.   □

Fix now a Christ's cube I and let $s_i(I), i = 1,...,M$, be all its sons. Let $a = (\langle w\rangle_{\mu,I}, \langle\sigma\rangle_{\mu,I})$, $b_i = (\langle w\rangle_{\mu,s_i(I)}, \langle\sigma\rangle_{\mu,s_i(I)})$, $i = 1,...,M$, be points–obviously–in Ω_Q, where Q temporarily means $[w]_{A_2}$. Consider $c_i(t) = a(1-t) + b_i t, 0 \leq t \leq 1$ and $q_i(t) := B_Q(c_i(t))$. We want to use Taylor's formula

(3.6) $$q_i(0) - q_i(1) = -q_i'(0) - \int_0^1 dx \int_0^x q_i''(t)\,dt.$$

Notice two things: Sublemma shows that $-q_i''(t) \geq 0$ always. Moreover, it shows that if $t \in [0, 1/2]$, then we have that the following qualitative estimate holds:

(3.7) $$-q_i''(t) \geq c(\langle w\rangle_{\mu,I}\langle\sigma\rangle_{\mu,I})^\alpha\left(\frac{(\langle w\rangle_{\mu,s_i(I)} - \langle w\rangle_{\mu,I})^2}{\langle w\rangle_{\mu,I}^2} + \frac{(\langle\sigma\rangle_{\mu,s_i(I)} - \langle\sigma\rangle_{\mu,I})^2}{\langle\sigma\rangle_{\mu,I}^2}\right)$$

This requires a small explanation. If we are on the segment $[a, b_i]$, then the first coordinate of such a point cannot be larger than $C\langle w\rangle_{\mu,I}$, where C depends only on doubling of μ (not w). This is obvious. The same is true for the second coordinate with the obvious change of w to σ. But there is no such type of estimate from below on this segment: the first coordinate cannot be smaller than $k\langle w\rangle_{\mu,I}$, but k may (and will) depend on the doubling of w (so ultimately on its $[w]_{A_2}$ norm. In fact, at the "right" endpoint of $[a, b_i]$. The first coordinate is $\langle w\rangle_{\mu,s_i(I)} \leq \int_I w\,d\mu/\mu(s_i(I)) \leq C\int_I w\,d\mu/\mu(I) = C\langle w\rangle_{\mu,I}$, with C only depending on the doubling of μ. But the estimate from below will involve the doubling

of w, which we must avoid. But if $t \in [0, 1/2]$, and we are on the "left half" of interval $[a, b_i]$ then obviously the first coordinate is $\geq \frac{1}{2}\langle w\rangle_{\mu,I}$ and the second coordinate is $\geq \frac{1}{2}\langle\sigma\rangle_{\mu,I}$.

We do not need to integrate $-q_i''(t)$ for all $t \in [0, 1]$ in (3.6). We can only use integration over $[0, 1/2]$ noticing that $-q_i''(t) \geq 0$ otherwise. Then the chain rule

$$q_i''(t) = (B_Q(c_i(t)))'' = (d^2 B_Q(c_i(t))(b_i - a), b_i - a),$$

(where $(\cdot, \cdot)$ means the usual scalar product in $\mathbb{R}^2$) immediately gives us (3.7) with constant c depending on the doubling of μ but *independent* of the doubling of w.

Next step is to add all (3.6), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$, and to notice that $\sum_{i=1}^{M} \frac{\mu(s_i(I))}{\mu(I)} q_i'(0) = \nabla B_Q(a) \sum_{i=1}^{M} \cdot (a - b_i)\frac{\mu(s_i(I))}{\mu(I)} = 0$, because by definition

$$a = \sum_{i=1}^{M} \frac{\mu(s_i(I))}{\mu(I)} b_i.$$

Notice that the addition of all (3.6), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$ gives us now (we take into account (3.7) and positivity of $-q_i''(t)$)

$$B_Q(a) - \sum_{i=1}^{M} \frac{\mu(s_i(I))}{\mu(I)} B_Q(b_i) \geq c\, c_1 \left(\langle w\rangle_{\mu,I}\langle\sigma\rangle_{\mu,I}\right)^\alpha \sum_{i=1}^{M} \left(\frac{(\langle w\rangle_{\mu,s_i(I)} - \langle w\rangle_{\mu,I})^2}{\langle w\rangle_{\mu,I}^2} + \frac{(\langle\sigma\rangle_{\mu,s_i(I)} - \langle\sigma\rangle_{\mu,I})^2}{\langle\sigma\rangle_{\mu,I}^2} \right).$$

We used here the doubling of μ again, by noticing that $\frac{\mu(s_i(I))}{\mu(I)} \geq c_1$ (recall that $s_i(I)$ and I are almost balls of comparable radii). We rewrite the previous inequality using our definition of $\Delta_I w, \Delta_I \sigma$ listed above as follows

$$\mu(I) B_Q(a) - \sum_{i=1}^{M} \mu(s_i(I)) B_Q(b_i) \geq c\, c_1 \left(\langle w\rangle_{\mu,I}\langle\sigma\rangle_{\mu,I}\right)^\alpha \left(\frac{(\Delta_I w)^2}{\langle w\rangle_{\mu,I}^2} + \frac{(\Delta_I \sigma)^2}{\langle\sigma\rangle_{\mu,I}^2} \right)\mu(I).$$

Notice that $B_Q(a) = \langle w\rangle_{\mu,I}\langle\sigma\rangle_{\mu,I}$. Now we iterate the above inequality and get for any of Christ's dyadic I's:

$$\sum_{J \subset I, J \in \mathscr{D}} \left(\langle w\rangle_{\mu,J}\langle\sigma\rangle_{\mu,J}\right)^\alpha \left(\frac{(\Delta_J w)^2}{\langle w\rangle_{\mu,J}^2} + \frac{(\Delta_J \sigma)^2}{\langle\sigma\rangle_{\mu,J}^2} \right)\mu(J) \leq C Q^\alpha \mu(I).$$

This is exactly the Carleson property of the measure $\{\mu_I\}$ indicated in our Lemma 3.2, with Carleson constant $C Q^\alpha$. The proof showed that C depended only on $\alpha \in (0, 1/2)$ and on the doubling constant of measure μ.

$\square$

Now, using this lemma, we start to estimate our S_L's and R_L's. For $S_L(\psi\sigma)$ we already had estimate (3.4).

To estimate $R_L(\phi w)$ let us denote by $\mathscr{P}_L$ maximal stopping intervals $K \in \mathscr{D}, K \subset L$, where the stopping criteria are 1) either $\frac{|\Delta_K w|}{\langle w\rangle_{\mu,K}} \geq \frac{1}{m+n+1}$, or $\frac{|\Delta_K \sigma|}{\langle\sigma\rangle_{\mu,K}} \geq \frac{1}{m+n+1}$, or 2) $g(K) = g(L) + m$.

Lemma 3.3. *If K is any stopping interval then*
(3.8)

$$\sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle\phi w\rangle_{\mu,I}| \frac{|\Delta_I w|}{\langle w\rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}} \leq 2e^\alpha (m+n+1)\langle|\phi|w\rangle_{\mu,K} \frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}} \sqrt{\mu_K} \langle w\rangle_{\mu,L}^{-\alpha/2} \langle\sigma\rangle_{\mu,L}^{-\alpha/2}.$$

Proof. If we stop by the first criterion, then

$$\sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle\phi w\rangle_{\mu,I}| \frac{|\Delta_I w|}{\langle w\rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}} \leq 2 \sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle\phi w\rangle_{\mu,I}| \mu(I) \frac{1}{\mu(K)} \frac{\mu(K)}{\sqrt{\mu(L)}} \leq 2 \langle|\phi|w\rangle_{\mu,K} \frac{\mu(K)}{\sqrt{\mu(L)}}$$

$$\leq 2(m+n+1)\langle|\phi|w\rangle_{\mu,K}\left(\frac{|\Delta_K w|}{\langle w\rangle_{\mu,K}}+\frac{|\Delta_K\sigma|}{\langle\sigma\rangle_{\mu,K}}\right)\frac{\mu(K)}{\sqrt{\mu(L)}}\leq 2(m+n+1)\langle|\phi|w\rangle_{\mu,K}\frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}}\sqrt{\mu_K}\langle w\rangle_{\mu,K}^{-\alpha/2}\langle\sigma\rangle_{\mu,K}^{-\alpha/2}.$$

Now replacing $\langle w\rangle_{\mu,K}^{-\alpha/2}\langle\sigma\rangle_{\mu,K}^{-\alpha/2}$ by $\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}$ does not grow the estimate by more than e^α as all pairs of son/father intervals larger than K and smaller than L will have there averages compared by constant at most $1\pm\frac{1}{m+n+1}$. And there are at most m such intervals between K and L.

If we stop by the second criterion, then K is one of I's, $g(I)=g(L)+m$, and

$$|\langle\phi w\rangle_{\mu,I}|\frac{|\Delta_I w|}{\langle w\rangle_{\mu,I}}\frac{\mu(I)}{\sqrt{\mu(L)}}\leq|\langle\phi w\rangle_{\mu,K}|\frac{\mu(K)}{\sqrt{\mu(L)}}\frac{|\Delta_K w|}{\langle w\rangle_{\mu,K}}\leq\langle|\phi|w\rangle_{\mu,K}\frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}}\sqrt{\mu_K}\langle w\rangle_{\mu,K}^{-\alpha/2}\langle\sigma\rangle_{\mu,K}^{-\alpha/2}.$$

Now we replace $\langle w\rangle_{\mu,K}^{-\alpha/2}\langle\sigma\rangle_{\mu,K}^{-\alpha/2}$ by $\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}$ as before.

$\square$

Now

$$R_L(\phi w)\leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\sum_{K\in\mathscr{P}_L}\langle|\phi|w\rangle_{\mu,K}\frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}}\sqrt{\mu_K}$$

$$\leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|w\rangle_{\mu,K}^2\frac{\mu(K)}{\mu(L)}\right)^{1/2}(\widetilde{\mu}(L))^{1/2},$$

where

$$\widetilde{\mu}_L=\sum_{K\in\mathscr{P}_L}\mu_K.$$

Notice that $\widetilde{\mu}_L$ form a Carleson measure with constant at most $C(m+1)Q^\alpha$.

Now we make a trick! We will estimate the right hand side as

$$R_L(\phi w)\leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|w\rangle_{\mu,K}^p\frac{\mu(K)}{\mu(L)}\right)^{1/p}(\widetilde{\mu}_L)^{1/2},$$

where $p=2-\frac{1}{m+n+1}$. In fact,

$$\left(\sum_{K\subset L, K\text{ is maximal}}\langle|\phi|w\rangle_{\mu,K}^2\frac{\mu(K)}{\mu(L)}\right)^{p/2}\leq\sum_{K\in\mathscr{P}_L}\langle|\phi|w\rangle_{\mu,K}^p\left(\frac{\mu(K)}{\mu(L)}\right)^{p/2}.$$

But if if $0\leq j\leq m$, then $(C^{-j})^{-\frac{1}{m+n+1}}\leq C$, and therefore in the formula above $\left(\frac{\mu(K)}{\mu(L)}\right)^{1-\frac{1}{2(m+n+1)}}\leq$ $C\frac{\mu(K)}{\mu(L)}$, and C depends only on the doubling constant of μ. So the trick is justified. Therefore, using Cauchy inequality, one gets

$$R_L(\phi w)\leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|^p w\rangle_{\mu,K}\langle w\rangle_{\mu,K}^{p-1}\frac{\mu(K)}{\mu(L)}\right)^{1/p}(\widetilde{\mu}_L)^{1/2}.$$

We can replace all $\langle w\rangle_{\mu,K}^{p-1}$ by $\langle w\rangle_{\mu,L}^{p-1}$ paying the price by constant. This is again because all intervals larger than K and smaller than L will have there averages compared by constant at most $1\pm\frac{1}{m+n+1}$. And there are at most m such intervals between K and L. Finally,

$$(3.9)\qquad R_L(\phi w)\leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|^p w\rangle_{\mu,K}\frac{\mu(K)}{\mu(L)}\right)^{1/p}\langle w\rangle_{\mu,L}^{1-\frac{1}{p}}(\widetilde{\mu}_L)^{1/2}$$

We need the standard notations: if ν is an arbitrary positive measure we denote

$$M_\nu f(x) := \sup_{r>0} \frac{1}{\nu(B(x,r))} \int_{B(x,r)} |f(x)| \, d\nu(x).$$

In particular M_w will stand for this maximal function with $d\nu = w(x)\,d\mu$.

From (3.9) we get

(3.10) $$R_L(\phi w) \le C(m+n+1)\langle w \rangle_{\mu,L}^{1-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2} \inf_L M_w(|\phi|^p)^{1/p} (\widetilde{\mu}_L)^{1/2}$$

Now
(3.11)
$$S_L(\psi\sigma)R_L(\phi w) \le C(m+n+1)\langle w \rangle_{\mu,L}^{1-\alpha/2} \langle \sigma \rangle_{\mu,L}^{1-\alpha/2} \frac{\inf_L M_w(|\phi|^p)^{1/p}}{\langle \sigma \rangle_{\mu,L}^{1/2}} (\widetilde{\mu}_L)^{1/2} \sqrt{\sum_{J \subset L\ldots} |(\psi\sigma, h_J^\sigma)|^2},$$

(3.12) $$R_L(\psi\sigma)R_L(\phi w) \le C(m+n+1)\langle w \rangle_{\mu,L}^{1-\alpha} \langle \sigma \rangle_{\mu,L}^{1-\alpha} \inf_L M_w(|\phi|^p)^{1/p} \inf_L M_\sigma(|\psi|^p)^{1/p} \widetilde{\mu}_L.$$

Now we use the Carleson property of $\{\widetilde{\mu}_L\}_{L \in \mathscr{D}}$. We need a simple folklore Lemma.

Lemma 3.4. *Let $\{\alpha_L\}_{L \in \mathscr{D}}$ define Carleson measure with intensity B related to dyadic lattice $\mathscr{D}$ on metric space X. Let F be a positive function on X. Then*

(3.13) $$\sum_L (\inf_L F)\,\alpha_L \le 2B \int_X F\,d\mu.$$

(3.14) $$\sum_L \frac{\inf_L F}{\langle \sigma \rangle_{\mu,L}} \alpha_L \le CB \int_X \frac{F}{\sigma}\,d\mu.$$

Now use (3.11). Then the estimate of $III \le \sum_L S_L(\psi\sigma)R_L(\phi w)$ will be reduced to estimating

$$(m+n+1)Q^{1-\alpha/2}\left(\sum_L \frac{\inf_L M_w(|\phi|^p)^{2/p}}{\langle \sigma \rangle_{\mu,L}} \widetilde{\mu}_L\right)^{1/2} \le (m+n+1)^2 Q\left(\int_{\mathbb{R}} (M_w(|\phi|^p))^{2/p} w\,d\mu\right)^{1/2}$$

$$\le \left(\frac{1}{2-p}\right)^{1/p}(m+n+1)^2 Q\left(\int_{\mathbb{R}} \phi^2 w\,d\mu\right)^{1/2} \le (m+n+1)^3 Q\left(\int_{\mathbb{R}} \phi^2 w\,d\mu\right)^{1/2}.$$

Here we used (3.14) and the usual estimates of maximal function M_μ in $L^q(\mu)$ when $q \approx 1$. Of course for II we use the symmetric reasoning.

Now IV: we use (3.12) first.

$$\sum_L S_L(\psi\sigma)R_L(\phi w) \le (m+n+1)Q^{1-\alpha}\sum_L \inf_L M_w(|\phi|^p)^{1/p} \inf_L M_\sigma(|\psi|^p)^{1/p} \widetilde{\mu}_L$$

$$\le C(m+n+1)^2 Q \int_{\mathbb{R}} (M_w(|\phi|^p))^{1/p} (M_\sigma(|\psi|^p))^{1/p} w^{1/2}\sigma^{1/2}\,d\mu$$

$$\le C(m+n+1)^2 Q \left(\int_{\mathbb{R}} (M_w(|\phi|^p))^{2/p} w\,d\mu\right)^{1/2} \left(\int_{\mathbb{R}} (M_\sigma(|\psi|^p))^{2/p} \sigma\,d\mu\right)^{1/2}$$

$$\le C(m+n+1)^4 Q \left(\int_{\mathbb{R}} \phi^2 w\,d\mu\right)^{1/2} \left(\int_{\mathbb{R}} \psi^2 \sigma\,d\mu\right)^{1/2}.$$

Here we used (3.13) and the usual estimates of maximal function M_μ in $L^{2/p}(\mu)$ when $p \approx 2$, $p < 2$.

References

[1] S. M. BUCKLEY, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc., **340** (1993), no. 1, p53–272.

[2] G. BEYLKIN, R. COIFMAN, V. ROKHLIN, *Fast wavelet transforms and numerical algorithms, I*, Comm. Pure and Appl. Math., **44** (1991), No. 2, 141–183.

[3] M. CHRIST, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601-628.

[4] T. HYTÖNEN, *The sharp weighted bound for general Calderon-Zygmund operators*, arXiv:1007.4330.

[5] T. HYTÖNEN, *Nonhomogeneous vector Tb theorem*, arXiv:0809.3097.

[6] T. HYTÖNEN, M. LACEY, M. C. REGUERA, E. SAWYER, A. VAGHARSHAKYAN, I. URIARTE-TUERO, em Weak and Strong type A_p Estimates for Calderãn-Zygmund Operators, arXiv:1006.2530.

[7] R. HUNT, B. MUCKENHOUPT, R. WHEEDEN, *Weighted norm inequalities for the conjugate function and the Hilbert transform*, Trans. Amer. Math. Soc., **176** (1973), pp. 227-251.

[8] T. HYTÖNEN, C. PÉREZ, S. TREIL, A. VOLBERG, A sharp estimated of weighted dyadic shifts that gives the proof of A_2 conjecture, Preprint, pp. 1–36, Sept. 2010.

[9] M. LACEY, S. PETERMICHL, M. RIGUERA, *Sharp A_2 inequality for Haar shift operators* arXiv:0906.1941.

[10] S. KONYAGIN, A. VOLBERG, *On measures with the doubling condition.* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **51** (1987), no. 3, 666–675; translation in Math. USSR-Izv. **30** (1988), no. 3, 629–638.

[11] A. LERNER *A pointwise estimate for local sharp maximal function with applications to singular integrals*, preprint, 2009.

[12] F. NAZAROV, A. VOLBERG, Bellman function, polynomial estimates of weighted dyadic shifts, and A_2 conjecture. Preprint, Nov. 2010, pp. 1–6.

[13] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Tb theorems on nonhomogeneous spaces*, Acta Math., **190** (2003), 151–239.

[14] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Preprint 2004, 1–14. Appeared in Math. Res. Lett. 15 (2008), no. 3, 583–597.

[15] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight $T1$ theorem for the Hilbert transform: the case of doubling measures*, Preprint 2004, 1–40.

[16] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures*, Preprint 2005, 1-33. Put into arXive in 2010.

[17] F. Nazarov, S. Treil and A. Volberg, *Bellman function in stochastic control and harmonic analysis.* Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 393–423, Oper. Theory Adv. Appl., **129**, Birkhauser, Basel, 2001.

[18] F. NAZAROV, S. TREIL, AND A. VOLBERG, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. of Amer. Math. Soc., **12**, (1999), no. 4, 909-928.

[19] S. PETERMICHL, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris, Sér. I Math., **330**, (2000), no. 6, pp. 455-460.

[20] S. PETERMICHL, A. VOLBERG, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J., **112** (2002), no. 2, pp. 281-305.

[21] S. PETERMICHL, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic.* Amer. J. Math. 129 (2007), no. 5, 1355–1375.

[22] S. PETERMICHL, *The sharp weighted bound for the Riesz transforms.* Proc. Amer. Math. Soc. 136 (2008), no. 4, 1237–1249.

[23] S. PETERMICHL, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol.* C. R. Acad. Sci. Paris S?r. I Math. 330 (2000), no. 6, 455–460.

[24] C. PÉREZ, S. TREIL, A. VOLBERG, *On A_2 conjecture and corona decomposition of weights*, arxiv1005.2630.

[25] C. PÉREZ, S. TREIL, A. VOLBERG, *A_2 conjecture: reduction to a bilinear embedding, splines and less smooth Calderón–Zygmund kernels*, Preprint, pp. 1–17, July 3, 2010.

[26] A. REZNIKOV, S. TREIL, A. VOLBERG, *A sharp estimate of weighted dyadic shifts of complexity 0 and 1*, Preprint, September 2010, pp. 1–9.

[27] A. REZNIKOV, A. VOLBERG, *Random "dyadic" lattice in geometrically doubling metric space and A_2 conjecture*, Preprint arXiv: 1. arXiv:1103.5246, pp. 1–10.

[28] V. VASYUNIN, A. VOLBERG, *Two weight inequality: the case study*, St. Petersburg Math. J., 2005?

[29] A. VOLBERG, *Calderón–Zygmund capacities and operators on non-homogeneous spaces*, CBMS conferences, AMS, vol. 100, 2003, 165pp.

[30] J. WITTWER, *A sharp estimate on the norm of the martingale transform.* Math. Res. Lett. 7 (2000), no. 1, 1–12.

Dept. of Mathematics, Univ. of Wisconsin, Madison

Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

A SIMPLE SHARP WEIGHTED ESTIMATE OF THE DYADIC SHIFTS ON METRIC SPACES WITH GEOMETRIC DOUBLING

FEDOR NAZAROV AND ALEXANDER VOLBERG

ABSTRACT. We give a short and simple polynomial estimate of the norm of weighted dyadic shift on metric space with geometric doubling, which is linear in the norm of the weight. Combined with the existence of special probability space of dyadic lattices built in A. Reznikov, A. Volberg, " Random "dyadic" lattice in geometrically doubling metric space and A_2 conjecture", arXiv:1103.5246, and with decomposition of Calderón–Zygmund operators to dyadic shifts from Hytonen's "The sharp weighted bound for general Calderon-Zygmund operators", arXiv:1007.4330 (and later T. Hytönen, C. Pérez, S. Treil, A. Volberg, "A sharp estimated of weighted dyadic shifts that gives the proof of A_2 conjecture", arXiv 1010.0755.), we will be able to have a linear (in the norm of weight) estimate of an arbitrary Calderón–Zygmund operator on a metric space with geometric doubling. This will be published separately.

1. INTRODUCTION

Recall that in [24] it was proved that

Theorem 1.1. *If T is an arbitrary operator with a Calderón–Zygmund kernel, then*

$$\|T\|_{L^2(wd\mu)\to L^{2,\infty}(w)} + \|T'\|_{L^2(w^{-1})\to L^{2,\infty}(w^{-1})} \le 2\|T\|_{L^2(wd\mu)\to L^2(wd\mu)}$$
$$\le C\left([w]_{A_2} + \|T\|_{L^2(wd\mu)\to L^{2,\infty}(w)} + \|T'\|_{L^2(w^{-1})\to L^{2,\infty}(w^{-1})}\right).$$

By T' we denote the adjoint operator. Here of course only the right inequality is interesting. And it is unexpected too. The weak and strong norm of any operator with a Calderón–Zygmund kernel turned out to be equivalent up to additive term $[w]_{A_2}$. In its turn, Theorem 1.1 was obtained in [24] as a corollary of a weighted $T1$ theorem–the Main Theorem of [24]. Again in its turn the Main Theorem (=*weighted $T1$ theorem*) in [24] is a consequence of a rather difficult two-weight $T1$ theorem of [16]. From Theorem 1.1 we obtained in [24] the following result which holds for any Calderón–Zygmund operator.

Theorem 1.2. $\|T\|_{L^2(wd\mu)\to L^2(wd\mu)} \le C\cdot[w]_{A_2}\log(1+[w]_{A_2})$.

By A_2 conjecture people understand the strengthening of this claim, where the logarithmic term is deleted, in other words, a linear (in weight's norm) estimate of arbitrary weighted Calderón–Zygmund operator. In [6] the A_2 conjecture was proved for Calderón–Zygmund operators having more than $2d$ smoothness in $\mathbb{R}^d$.

A bit later a preprint [4] of Tuomas Hytönen has appeared, the A_2 conjecture is fully proved there. It is based on the Main Theorem (=*weighted $T1$ theorem*) in [24] of Pérez–Treil–Volberg. Both [24] and [4] are neither short nor easy.

Notice that the scheme of the proof in [4] goes like that: given a Calderón–Zygmund operator T and a $w \in A_2$ weight, one first uses the Main Theorem (=*weighted $T1$ theorem*) of [24], which

says that to prove a linear estimate for $\|Tf\|_w$ it is enough to prove it uniformly only for special "characteristic functions of cubes" type functions f (see the true statement in [24]). Notice that all the cubes must be checked. The second (very beautiful) part of the proof is to decompose T into "a convex combinations" of dyadic shifts, the new idea is used here that grew out of random lattices approach in non-homogeneous Harmonic Analysis theory of Nazarov–Treil–Volberg, see [13], [15], [16], [29]. Now it is enough to check the right estimate for each dyadic shift applied to each "characteristic function of cube". A very annoying difficulty appears: the shift is with respect to a certain dyadic lattice, but the cube in question is arbitrary and a priori does not belong to this lattice. This creates serious technical difficulties, which can be (and were) avoided in [8].

The direct proof of A_2 conjecture (without going through [24]) was given in [8], and it was based on two ingredients: 1) a formula for decomposing an arbitrary Calderón–Zygmund operators into (generalized) dyadic shifts by the averaging trick, see [4], 2) on a polynomial in complexity and linear in weight estimate of the norm of a dyadic shift.

The latter was quite complicated and was based on modification of the argument in Lacey–Petermichl-Reguera [9] and on the use of [14] with its careful reexamination. The former–as we already mentioned–was rooted in the works on non-homogeneous Harmonic Analysis, like e. g. [13]–[15], [16], [29], but with a new twist, which appeared first in Hytönen's [4] and was simplified in Hytönen–Pérez–Treil–Volberg's [8].

The averaging trick was a development of the bootstrapping argument used by Nazarov–Treil–Volberg, where they exploited the fact that the bad part of a function can be made arbitrarily small. Using the original Nazarov–Treil–Volberg averaging trick would add an extra factor depending on $[w]_{A_2}$ to the estimate, so a new idea was necessary. The new observation in [4] was that as soon as the probability of a "bad" cube is less than 1, it is possible to completely ignore the bad cubes (at least in the situation where they cause troubles).

2. Shifts of complexity m, n

Let X be a compact geometrically doubling metric space. Let μ be a doubling measure on X, which exists by [10]. Let $\mathscr{D}$ be a finite "dyadic" lattice on X. Namely, $\mathscr{D}$ consists of disjoint partition $\mathscr{E}_N$ to pieces Q_N^j of size $\asymp \delta^N$, then there is a partition $\mathscr{E}_{N-1}$ to pieces Q_{N-1}^i of size δ^{N-1}, each Q_{N-1}^i consists of boundedly many (at most $M_{N-1}(i) \le M$) pieces of partition $\mathscr{E}_N$, et cetera... , we have $\mathscr{E}_k, k = N, N-1, ..., 0$, and each Q_k^i is an almost ball: if $\ell(Q), Q = Q_k^i$, denote its diameter, then there is a ball of radius $c\ell(Q), c > 0$, inside Q, and c does not depend on N, k, i.

The existence of such lattices with all constants depending only on geometric doubling of X was proved by Christ [3].

We relate to $\mathscr{E}_k$ the projection operator $\mathbb{E}_k$ on $L^2(\mu)$: $\mathbb{E}_k f = \sum_j \langle f \rangle_{Q_k^j, \mu} \chi_{Q_k^j}$. We also consider the martingale difference operator $\Delta_k f = E_k f - E_{k-1} f$. Notice that it can be written as

$$\Delta_k f = \sum_i \sum_{j=1}^{M_{k-1}(Q_{k-1}^i)} (f, h_{Q_{k-1}^i}^j)_\mu h_{Q_{k-1}^i}^j ,$$

where denoting $Q = Q_{k-1}^i$ we notice that $M_{k-1}(Q) \le M$, and h_Q^j, are functions supported on Q, orthogonal to constants in $L^2(\mu)$, orthogonal to each other in $L^2(\mu)$, constant on each $Q_k^i \subset Q, Q_k^i \in \mathscr{E}_k$ (such Q_k^i are called the **sons** of Q_{k-1}^j), and having the following bound

(2.1) $$\|h_Q^j\|_{L^\infty(\mu)} \le \frac{C}{\mu(Q)^{1/2}} .$$

Definition. We call such $h_{Q_k}^j$ Haar functions. If $L \in \mathscr{E}_m, J \in \mathscr{E}_{m+n}$ we say $g(J) = g(L) + n$.

We always use $\ell(S)$ to denote the diameter of a set S. Christ's cubes will be denoted Q, I, J, L, K, may be with indices.

We call by $\mathbb{S}_{m,n}$ the operator given by the kernel

$$f \to \sum_{L \in \mathscr{D}} \int_L a_L(x,y) f(y) dy \,,$$

where

$$a_L(x,y) = \sum_{\substack{I \subset L, J \subset L \\ g(I) = g(L) + m, g(J) = g(L) + n}} c_{L,I,J} h_J^j(x) h_I^i(y) \,,$$

where h_I^i, h_J^j are Haar functions normalized in $L^2(d\mu)$ and satisfying (2.1), and $|c_{L,I,J}| \le \frac{\sqrt{\mu(I)}\sqrt{\mu(J)}}{\mu(L)}$. Often we will skip superscripts i, j.

We are interested in sufficiently good estimate of

$$\|\mathbb{S}_{m,n}\|_w := \|\mathbb{S}_{m,n} : L^2(w\,d\mu) \to L^2(w\,d\mu)\| \,,$$

where $w \in A_2$. For such w we put $\sigma = w^{-1}$ and

$$[w]_{A_2} := \sup_I \langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I} < \infty \,,$$

and call it the norm of w (it is not a norm).

In recent paper [8] the following theorem was proved (another proof, using the Bellman function technique, was given recently in [12])

Theorem 2.1.

(2.2) $$\|\mathbb{S}_{m,n}\|_w \le C (m+n+1)^a [w]_{A_2} \,.$$

In [8] $a = 3$. Looks like here we have the same numerical value. But for its main application in [8]: the proof of A_2 conjecture, the value of a (if finite) is not important. The proof was hard and combinatorial, it was based on the ideas of [9], where such an estimate was proved with exponential dependence on $m + n$. We propose here a simple proof based on Bellman function technique. This technique was tried successfully for shifts of low complexity, first in [30], [20] then in [21], [22], and recently in preprint [26], which gives a simple unified proof of results in [30], [21], [22]. As the reader will see one needs a couple of new tricks to achieve this fulfillment. The proof below is a direct and simple readjustment of the proof in [12], where it has been carried out if $X = \mathbb{R}$, $\mu = dx$.

Remark. The reasoning below is in $\mathbb{R}$. But one can modify it without any efforts to any $\mathbb{R}^d$. Moreover, in [27] the probability space of Christ's type dyadic lattices is built on any compact metric space with the property of geometric doubling (every ball contains at most a fixed number of disjoint balls of half a radius), which allows to extend the sharp bound of Calderón–Zygmund operators into metric space setting by repeating the averaging trick that reduces everything to the case of dyadic shift on the metric space, and then using this preprint to give a polynomial in complexity and linear in weight estimate for any shift.

3. THE HEART OF THE MATTER: A REDUCTION TO BILINEAR EMBEDDING ESTIMATE

To prove Theorem 2.1 we need the following decomposition:

Lemma 3.1.
$$h_I^j = \alpha_I^j h_I^{w,j} + \beta_I^j \chi_I,$$

where

1) $|\alpha_I^j| \leq \sqrt{\langle w \rangle_{\mu,I}},$

2) $|\beta_I^j| \leq \frac{(h_I^{w,j},w)_\mu}{w(I)},$ *where* $w(I) := \int_I w \, d\mu,$

3) $\{h_I^{w,j}\}_I$ *is supported on I, orthogonal to constants in* $L^2(w \, d\mu),$

4) $h_I^{w,j}$ *assumes on each son $s(I)$ a constant values,*

5) $\|h_I^{w,j}\|_{L^2(\mu)} = 1.$

Definition. Let
$$\Delta_I w := \sum_{\text{sons of } I} |\langle w \rangle_{\mu,s(I)} - \langle w \rangle_{\mu,I}|.$$

Remark. There are many Δ's in this paper, but the reader should notice that Δ_k is an operator sending functions from $L^2(\mu)$ to such functions with extra properties of being constants on each element of $\mathcal{E}_k$ and also being orthogonal in $L^2(\mu)$ to χ_L for any element L of $\mathcal{E}_{k-1}$. On the other hand, $\Delta_I w$ (as well as $\Delta_J \sigma$) is a non-negative number.

It is a easy to see (by a small linear algebra reasoning) that the doubling property of measure μ implies

(3.1) $$|(h_I^{w,j},w)_\mu| \leq C \cdot (\Delta_I w) \sqrt{\mu(I)}.$$

Here C depends only on the doubling constant of measure μ. In other words, taking into account that $\Delta_k w = \sum_{I \in \mathcal{E}_{k-1}} \sum_{j=1}^{M_{k-1}(I)} (w,h_I^{w,j})_\mu h_I^{w,j}$ and given $I \in \mathcal{E}_{k-1}$, we can rewrite (3.1) as follows

(3.2) $$\chi_I |\Delta_k w| \leq C \cdot (\Delta_I w),$$

where C again depends only on the doubling constant of μ.

Therefore, the property 2) above can be rewritten as

2') $|\beta_I^j| \leq C \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{1}{\sqrt{\mu(I)}}.$

Fix $\phi \in L^2(w \, d\mu), \psi \in L^2(\sigma)$. We need to prove

(3.3) $$|(\mathbb{S}_{m,n} \phi w, \psi \sigma)| \leq C(n+m+1)^a \|\phi\|_w \|\psi\|_\sigma.$$

We estimate $(\mathbb{S}_{m,n} \phi w, \psi \sigma)$ as (we skip superscripts j,i and write just h_J, h_I):

$$|\sum_L \sum_{I,J} c_{L,I,J} (\phi w, h_I)_\mu (\psi \sigma, h_J)_\mu| \leq$$

$$\sum_L \sum_{I,J} |c_{L,I,J} (\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} (\psi \sigma, h_J^\sigma)_\mu| \sqrt{\langle \sigma \rangle_{\mu,J}} | +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \phi w \rangle_{\mu,I} \frac{\Delta_I w}{\langle w \rangle_{\mu,I}} (\psi \sigma, h_J^\sigma)_\mu \sqrt{\langle \sigma \rangle_{\mu,J}} \sqrt{\mu(I)} | +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \psi \sigma \rangle_{\mu,J} \frac{\Delta_J \sigma}{\langle \sigma \rangle_{\mu,J}} (\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} \sqrt{\mu(J)} | +$$

$$\sum_L \sum_{I,J} |c_{L,I,J} \langle \phi w \rangle_{\mu,I} \langle \psi \sigma \rangle_{\mu,J} \frac{\Delta_I w}{\langle w \rangle_{\mu,I}} \frac{\Delta_J \sigma}{\langle \sigma \rangle_{\mu,J}} \sqrt{\mu(I)} \sqrt{\mu(J)}| =: I + II + III + IV.$$

We can notice that because we have $|c_{L,I,J}| \le \frac{\sqrt{\mu(I)}\sqrt{\mu(J)}}{\mu(L)}$, each sum inside L can be estimated by a perfect product of S and R terms, where

$$R_L(\phi w) := \sum_{I \subset L \ldots} \langle \phi w \rangle_{\mu,I} \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}}$$

$$S_L(\phi w) := \sum_{I \subset L \ldots} (\phi w, h_I^w)_\mu \sqrt{\langle w \rangle_{\mu,I}} \frac{\sqrt{\mu(I)}}{\sqrt{\mu(L)}}$$

and the corresponding terms for $\psi \sigma$. So we have

$$I \le \sum_L S_L(\phi w) S_L(\psi \sigma), \, II \le \sum_L S_L(\phi w) R_L(\psi \sigma),$$

$$III \le \sum_L R_L(\phi w) S_L(\psi \sigma), \, IV \le \sum_L R_L(\phi w) R_L(\psi \sigma).$$

Now

(3.4) $\qquad S_L(\phi w) \le \sqrt{\sum_{I \subset L \ldots} |(\phi w, h_I^w)_\mu|^2} \sqrt{\langle w \rangle_{\mu,L}}, \, S_L(\psi \sigma) \le \sqrt{\sum_{J \subset L \ldots} |(\psi \sigma, h_J^\sigma)|^2} \sqrt{\langle \sigma \rangle_{\mu,L}}$

Therefore,

(3.5) $$I \le C[w]_{A_2}^{1/2} \|\phi\|_w \|\psi\|_\sigma.$$

Terms II, III are symmetric, so consider III. Using Bellman function $(xy)^\alpha$ one can prove now

Lemma 3.2. *The sequence*

$$\mu_I := \langle w \rangle_{\mu,I}^\alpha \langle \sigma \rangle_{\mu,I}^\alpha \left(\frac{|\Delta_I w|^2}{\langle w \rangle_{\mu,I}^2} + \frac{|\Delta_I \sigma|^2}{\langle \sigma \rangle_{\mu,I}^2} \right) \mu(I)$$

form a Carleson measure with Carleson constant at most $c_\alpha Q^\alpha$, where $Q := [w]_{A_2}$ for any $\alpha \in (0, 1/2)$. Here c_α depends only on α and the doubling constant of μ (and is independent of the doubling constants of $w\, d\mu$, $\sigma\, d\mu$).

Proof. We need a very simple

Sublemma. Let $Q > 1, 0 < \alpha < \frac{1}{2}$. In domain $\Omega_Q := \{(x,y) : X > o, y > 0, 1 < xy \le Q$ function $B_Q(x,y) := x^\alpha y^\alpha\}$ satisfies the following estimate of its Hessian matrix (of its second differential form, actually)

$$-d^2 B_Q(x,y) \ge \alpha(1 - 2\alpha) x^\alpha y^\alpha \left(\frac{(dx)^2}{x^2} + \frac{(dy)^2}{y^2} \right).$$

The form $-d^2 B_Q(x,y) \ge 0$ everywhere in $x > 0, y > 0$. Also obviously $0 \le B_Q(x,y) \le Q^\alpha$ in Ω_Q.

Proof. Direct calculation. $\qquad \square$

Fix now a Christ's cube I and let $s_i(I), i = 1, ..., M$, be all its sons. Let $a = (\langle w \rangle_{\mu,I}, \langle \sigma \rangle_{\mu,I})$, $b_i = (\langle w \rangle_{\mu,s_i(I)}, \langle \sigma \rangle_{\mu,s_i(I)})$, $i = 1, \ldots, M$, be points–obviously–in Ω_Q, where Q temporarily means $[w]_{A_2}$. Consider $c_i(t) = a(1-t) + b_i t, 0 \le t \le 1$ and $q_i(t) := B_Q(c_i(t))$. We want to use Taylor's formula

(3.6) $$q_i(0) - q_i(1) = -q_i'(0) - \int_0^1 dx \int_0^x q_i''(t)\, dt.$$

Notice two things: Sublemma shows that $-q_i''(t) \geq 0$ always. Moreover, it shows that if $t \in [0, 1/2]$, then we have that the following qualitative estimate holds:

$$(3.7) \qquad -q_i''(t) \geq c\left(\langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I}\right)^\alpha \left(\frac{(\langle w \rangle_{\mu,s_i(I)} - \langle w \rangle_{\mu,I})^2}{\langle w \rangle_{\mu,I}^2} + \frac{(\langle \sigma \rangle_{\mu,s_i(I)} - \langle \sigma \rangle_{\mu,I})^2}{\langle \sigma \rangle_{\mu,I}^2} \right)$$

This requires a small explanation. If we are on the segment $[a, b_i]$, then the first coordinate of such a point cannot be larger than $C \langle w \rangle_{\mu,I}$, where C depends only on doubling of μ (not w). This is obvious. The same is true for the second coordinate with the obvious change of w to σ. But there is no such type of estimate from below on this segment: the first coordinate cannot be smaller than $k \langle w \rangle_{\mu,I}$, but k may (and will) depend on the doubling of w (so ultimately on its $[w]_{A_2}$ norm. In fact, at the "right" endpoint of $[a, b_i]$. The first coordinate is $\langle w \rangle_{\mu,s_i(I)} \leq \int_I w \, d\mu / \mu(s_i(I)) \leq C \int_I w \, d\mu / \mu(I) = C \langle w \rangle_{\mu,I}$, with C only depending on the doubling of μ. But the estimate from below will involve the doubling of w, which we must avoid. But if $t \in [0, 1/2]$, and we are on the "left half" of interval $[a, b_i]$ then obviously the first coordinate is $\geq \frac{1}{2} \langle w \rangle_{\mu,I}$ and the second coordinate is $\geq \frac{1}{2} \langle \sigma \rangle_{\mu,I}$.

We do not need to integrate $-q_i''(t)$ for all $t \in [0, 1]$ in (3.6). We can only use integration over $[0, 1/2]$ noticing that $-q_i''(t) \geq 0$ otherwise. Then the chain rule

$$q_i''(t) = (B_Q(c_i(t)))'' = (d^2 B_Q(c_i(t))(b_i - a), b_i - a),$$

(where $(\cdot, \cdot)$ means the usual scalar product in $\mathbb{R}^2$) immediately gives us (3.7) with constant c depending on the doubling of μ but *independent* of the doubling of w.

Next step is to add all (3.6), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$, and to notice that $\sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} q_i'(0) = \nabla B_Q(a) \sum_{i=1}^M \cdot (a - b_i) \frac{\mu(s_i(I))}{\mu(I)} = 0$, because by definition

$$a = \sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} b_i .$$

Notice that the addition of all (3.6), with convex coefficients $\frac{\mu(s_i(I))}{\mu(I)}$ gives us now (we take into account (3.7) and positivity of $-q_i''(t)$)

$$B_Q(a) - \sum_{i=1}^M \frac{\mu(s_i(I))}{\mu(I)} B_Q(b_i) \geq c\, c_1 \left(\langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I}\right)^\alpha \sum_{i=1}^M \left(\frac{(\langle w \rangle_{\mu,s_i(I)} - \langle w \rangle_{\mu,I})^2}{\langle w \rangle_{\mu,I}^2} + \frac{(\langle \sigma \rangle_{\mu,s_i(I)} - \langle \sigma \rangle_{\mu,I})^2}{\langle \sigma \rangle_{\mu,I}^2} \right).$$

We used here the doubling of μ again, by noticing that $\frac{\mu(s_i(I))}{\mu(I)} \geq c_1$ (recall that $s_i(I)$ and I are almost balls of comparable radii). We rewrite the previous inequality using our definition of $\Delta_I w, \Delta_I \sigma$ listed above as follows

$$\mu(I) B_Q(a) - \sum_{i=1}^M \mu(s_i(I)) B_Q(b_i) \geq c\, c_1 \left(\langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I}\right)^\alpha \left(\frac{(\Delta_I w)^2}{\langle w \rangle_{\mu,I}^2} + \frac{(\Delta_I \sigma)^2}{\langle \sigma \rangle_{\mu,I}^2} \right) \mu(I).$$

Notice that $B_Q(a) = \langle w \rangle_{\mu,I} \langle \sigma \rangle_{\mu,I}$. Now we iterate the above inequality and get for any of Christ's dyadic I's:

$$\sum_{J \subset I, J \in \mathscr{D}} \left(\langle w \rangle_{\mu,J} \langle \sigma \rangle_{\mu,J}\right)^\alpha \left(\frac{(\Delta_J w)^2}{\langle w \rangle_{\mu,J}^2} + \frac{(\Delta_J \sigma)^2}{\langle \sigma \rangle_{\mu,J}^2} \right) \mu(J) \leq C Q^\alpha \mu(I).$$

This is exactly the Carleson property of the measure $\{\mu_I\}$ indicated in our Lemma 3.2, with Carleson constant $C Q^\alpha$. The proof showed that C depended only on $\alpha \in (0, 1/2)$ and on the doubling constant of measure μ.

$\square$

Now, using this lemma, we start to estimate our S_L's and R_L's. For $S_L(\psi\sigma)$ we already had estimate (3.4).

To estimate $R_L(\phi w)$ let us denote by $\mathscr{P}_L$ maximal stopping intervals $K \in \mathscr{D}, K \subset L$, where the stopping criteria are 1) either $\frac{|\Delta_K w|}{\langle w \rangle_{\mu,K}} \geq \frac{1}{m+n+1}$, or $\frac{|\Delta_K \sigma|}{\langle \sigma \rangle_{\mu,K}} \geq \frac{1}{m+n+1}$, or 2) $g(K) = g(L) + m$.

Lemma 3.3. *If K is any stopping interval then*
(3.8)
$$\sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle \phi w \rangle_{\mu,I}| \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}} \leq 2e^\alpha (m+n+1)\langle |\phi|w \rangle_{\mu,K} \frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}} \sqrt{\mu_K} \langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2}.$$

Proof. If we stop by the first criterion, then

$$\sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle \phi w \rangle_{\mu,I}| \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}} \leq 2 \sum_{I \subset K, \ell(I)=2^{-m}\ell(L)} |\langle \phi w \rangle_{\mu,I}| \mu(I) \frac{1}{\mu(K)} \frac{\mu(K)}{\sqrt{\mu(L)}} \leq 2 \langle |\phi|w \rangle_{\mu,K} \frac{\mu(K)}{\sqrt{\mu(L)}}$$

$$\leq 2(m+n+1)\langle |\phi|w \rangle_{\mu,K} \left(\frac{|\Delta_K w|}{\langle w \rangle_{\mu,K}} + \frac{|\Delta_K \sigma|}{\langle \sigma \rangle_{\mu,K}} \right) \frac{\mu(K)}{\sqrt{\mu(L)}} \leq 2(m+n+1)\langle |\phi|w \rangle_{\mu,K} \frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}} \sqrt{\mu_K} \langle w \rangle_{\mu,K}^{-\alpha/2} \langle \sigma \rangle_{\mu,K}^{-\alpha/2}.$$

Now replacing $\langle w \rangle_{\mu,K}^{-\alpha/2} \langle \sigma \rangle_{\mu,K}^{-\alpha/2}$ by $\langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2}$ does not grow the estimate by more than e^α as all pairs of son/father intervals larger than K and smaller than L will have there averages compared by constant at most $1 \pm \frac{1}{m+n+1}$. And there are at most m such intervals between K and L.

If we stop by the second criterion, then K is one of I's, $g(I) = g(L) + m$, and

$$|\langle \phi w \rangle_{\mu,I}| \frac{|\Delta_I w|}{\langle w \rangle_{\mu,I}} \frac{\mu(I)}{\sqrt{\mu(L)}} \leq |\langle \phi w \rangle_{\mu,K}| \frac{\mu(K)}{\sqrt{\mu(L)}} \frac{|\Delta_K w|}{\langle w \rangle_{\mu,K}} \leq \langle |\phi|w \rangle_{\mu,K} \frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}} \sqrt{\mu_K} \langle w \rangle_{\mu,K}^{-\alpha/2} \langle \sigma \rangle_{\mu,K}^{-\alpha/2}.$$

Now we replace $\langle w \rangle_{\mu,K}^{-\alpha/2} \langle \sigma \rangle_{\mu,K}^{-\alpha/2}$ by $\langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2}$ as before. $\qquad \square$

Now

$$R_L(\phi w) \leq C(m+n+1)\langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2} \sum_{K \in \mathscr{P}_L} \langle |\phi|w \rangle_{\mu,K} \frac{\sqrt{\mu(K)}}{\sqrt{\mu(L)}} \sqrt{\mu_K}$$

$$\leq C(m+n+1)\langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2} \left(\sum_{K \in \mathscr{P}_L} \langle |\phi|w \rangle_{\mu,K}^2 \frac{\mu(K)}{\mu(L)} \right)^{1/2} (\widetilde{\mu}(L))^{1/2},$$

where

$$\widetilde{\mu}_L = \sum_{K \in \mathscr{P}_L} \mu_K.$$

Notice that $\widetilde{\mu}_L$ form a Carleson measure with constant at most $C(m+1)Q^\alpha$.

Now we make a trick! We will estimate the right hand side as

$$R_L(\phi w) \leq C(m+n+1)\langle w \rangle_{\mu,L}^{-\alpha/2} \langle \sigma \rangle_{\mu,L}^{-\alpha/2} \left(\sum_{K \in \mathscr{P}_L} \langle |\phi|w \rangle_{\mu,K}^p \frac{\mu(K)}{\mu(L)} \right)^{1/p} (\widetilde{\mu}_L)^{1/2},$$

where $p = 2 - \frac{1}{m+n+1}$. In fact,

$$\left(\sum_{K \subset L, K \text{ is maximal}} \langle |\phi|w \rangle_{\mu,K}^2 \frac{\mu(K)}{\mu(L)} \right)^{p/2} \leq \sum_{K \in \mathscr{P}_L} \langle |\phi|w \rangle_{\mu,K}^p \left(\frac{\mu(K)}{\mu(L)} \right)^{p/2}.$$

But if if $0 \leq j \leq m$, then $(C^{-j})^{-\frac{1}{m+n+1}} \leq C$, and therefore in the formula above $\left(\frac{\mu(K)}{\mu(L)}\right)^{1-\frac{1}{2(m+n+1)}} \leq C\frac{\mu(K)}{\mu(L)}$, and C depends only on the doubling constant of μ. So the trick is justified. Therefore, using Cauchy inequality, one gets

$$R_L(\phi w) \leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|^p w\rangle_{\mu,K}\langle w\rangle_{\mu,K}^{p-1}\frac{\mu(K)}{\mu(L)}\right)^{1/p}(\widetilde{\mu}_L)^{1/2}.$$

We can replace all $\langle w\rangle_{\mu,K}^{p-1}$ by $\langle w\rangle_{\mu,L}^{p-1}$ paying the price by constant. This is again because all intervals larger than K and smaller than L will have there averages compared by constant at most $1\pm\frac{1}{m+n+1}$. And there are at most m such intervals between K and L. Finally,

$$(3.9)\qquad R_L(\phi w) \leq C(m+n+1)\langle w\rangle_{\mu,L}^{-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\left(\sum_{K\in\mathscr{P}_L}\langle|\phi|^p w\rangle_{\mu,K}\frac{\mu(K)}{\mu(L)}\right)^{1/p}\langle w\rangle_{\mu,L}^{1-\frac{1}{p}}(\widetilde{\mu}_L)^{1/2}$$

We need the standard notations: if ν is an arbitrary positive measure we denote

$$M_\nu f(x) := \sup_{r>0}\frac{1}{\nu(B(x,r))}\int_{B(x,r)}|f(x)|\,d\nu(x).$$

In particular M_w will stand for this maximal function with $d\nu = w(x)\,d\mu$.

From (3.9) we get

$$(3.10)\qquad R_L(\phi w) \leq C(m+n+1)\langle w\rangle_{\mu,L}^{1-\alpha/2}\langle\sigma\rangle_{\mu,L}^{-\alpha/2}\inf_L M_w(|\phi|^p)^{1/p}(\widetilde{\mu}_L)^{1/2}$$

Now
(3.11)

$$S_L(\psi\sigma)R_L(\phi w) \leq C(m+n+1)\langle w\rangle_{\mu,L}^{1-\alpha/2}\langle\sigma\rangle_{\mu,L}^{1-\alpha/2}\frac{\inf_L M_w(|\phi|^p)^{1/p}}{\langle\sigma\rangle_{\mu,L}^{1/2}}(\widetilde{\mu}_L)^{1/2}\sqrt{\sum_{J\subset L\ldots}|(\psi\sigma,h_J^\sigma)|^2},$$

$$(3.12)\qquad R_L(\psi\sigma)R_L(\phi w) \leq C(m+n+1)\langle w\rangle_{\mu,L}^{1-\alpha}\langle\sigma\rangle_{\mu,L}^{1-\alpha}\inf_L M_w(|\phi|^p)^{1/p}\inf_L M_\sigma(|\psi|^p)^{1/p}\widetilde{\mu}_L.$$

Now we use the Carleson property of $\{\widetilde{\mu}_L\}_{L\in\mathscr{D}}$. We need a simple folklore Lemma.

Lemma 3.4. *Let $\{\alpha_L\}_{L\in\mathscr{D}}$ define Carleson measure with intensity B related to dyadic lattice $\mathscr{D}$ on metric space X. Let F be a positive function on X. Then*

$$(3.13)\qquad\qquad\sum_L(\inf_L F)\,\alpha_L \leq 2B\int_X F\,d\mu.$$

$$(3.14)\qquad\qquad\sum_L\frac{\inf_L F}{\langle\sigma\rangle_{\mu,L}}\alpha_L \leq CB\int_X\frac{F}{\sigma}d\mu.$$

Now use (3.11). Then the estimate of $III \leq \sum_L S_L(\psi\sigma)R_L(\phi w)$ will be reduced to estimating

$$(m+n+1)Q^{1-\alpha/2}\left(\sum_L\frac{\inf_L M_w(|\phi|^p)^{2/p}}{\langle\sigma\rangle_{\mu,L}}\widetilde{\mu}_L\right)^{1/2} \leq (m+n+1)^2Q\left(\int_{\mathbb{R}}(M_w(|\phi|^p))^{2/p}w\,d\mu\right)^{1/2}$$

$$\leq(\frac{1}{2-p})^{1/p}(m+n+1)^2Q\left(\int_{\mathbb{R}}\phi^2 w\,d\mu\right)^{1/2} \leq (m+n+1)^3Q\left(\int_{\mathbb{R}}\phi^2 w\,d\mu\right)^{1/2}.$$

Here we used (3.14) and the usual estimates of maximal function M_μ in $L^q(\mu)$ when $q\approx 1$. Of course for II we use the symmetric reasoning.

Now IV: we use (3.12) first.

$$\sum_L S_L(\psi\sigma)R_L(\phi w) \leq (m+n+1)Q^{1-\alpha}\sum_L \inf_L M_w(|\phi|^p)^{1/p}\inf_L M_\sigma(|\psi|^p)^{1/p}\widetilde{\mu}_L$$

$$\leq C(m+n+1)^2 Q\int_{\mathbb{R}}(M_w(|\phi|^p))^{1/p}(M_\sigma(|\psi|^p))^{1/p}w^{1/2}\sigma^{1/2}d\mu$$

$$\leq C(m+n+1)^2 Q\left(\int_{\mathbb{R}}(M_w(|\phi|^p))^{2/p}wd\mu\right)^{1/2}\left(\int_{\mathbb{R}}(M_\sigma(|\psi|^p))^{2/p}\sigma d\mu\right)^{1/2}$$

$$\leq C(m+n+1)^4 Q\left(\int_{\mathbb{R}}\phi^2 wd\mu\right)^{1/2}\left(\int_{\mathbb{R}}\psi^2\sigma d\mu\right)^{1/2}.$$

Here we used (3.13) and the usual estimates of maximal function M_μ in $L^{2/p}(\mu)$ when $p\approx 2$, $p<2$.

References

[1] S. M. BUCKLEY, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc., **340** (1993), no. 1, p53–272.

[2] G. BEYLKIN, R. COIFMAN, V. ROKHLIN, *Fast wavelet transforms and numerical algorithms, I*, Comm. Pure and Appl. Math., **44** (1991), No. 2, 141–183.

[3] M. CHRIST, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, Colloq. Math. **60/61** (1990), no. 2, 601-628.

[4] T. HYTÖNEN, *The sharp weighted bound for general Calderon-Zygmund operators*, arXiv:1007.4330.

[5] T. HYTÖNEN, *Nonhomogeneous vector Tb theorem*, arXiv:0809.3097.

[6] T. HYTÖNEN, M. LACEY, M. C. REGUERA, E. SAWYER, A. VAGHARSHAKYAN, I. URIARTE-TUERO, em Weak and Strong type A_p Estimates for Calderãn-Zygmund Operators, arXiv:1006.2530.

[7] R. HUNT, B. MUCKENHOUPT, R. WHEEDEN, *Weighted norm inequalities for the conjugate function and the Hilbert transform*, Trans. Amer. Math. Soc., **176** (1973), pp. 227-251.

[8] T. HYTÖNEN, C. PÉREZ, S. TREIL, A. VOLBERG, A sharp estimated of weighted dyadic shifts that gives the proof of A_2 conjecture, Preprint Sept. 2010, Preprint arXiv 1010.0755.

[9] M. LACEY, S. PETERMICHL, M. RIGUERA, *Sharp A_2 inequality for Haar shift operators* arXiv:0906.1941.

[10] S. KONYAGIN, A. VOLBERG, *On measures with the doubling condition.* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **51** (1987), no. 3, 666–675; translation in Math. USSR-Izv. **30** (1988), no. 3, 629–638.

[11] A. LERNER *A pointwise estimate for local sharp maximal function with applications to singular integrals*, preprint, 2009.

[12] F. NAZAROV, A. VOLBERG, Bellman function, polynomial estimates of weighted dyadic shifts, and A_2 conjecture. Preprint, Nov. 2010, pp. 1–6.

[13] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Tb theorems on nonhomogeneous spaces*, Acta Math., **190** (2003), 151–239.

[14] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Preprint 2004, 1–14. Appeared in Math. Res. Lett. 15 (2008), no. 3, 583–597.

[15] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight $T1$ theorem for the Hilbert transform: the case of doubling measures*, Preprint 2004, 1–40.

[16] F. NAZAROV, S. TREIL, AND A. VOLBERG, *Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures*, Preprint 2005, 1-33. Put into arXive in 2010.

[17] F. Nazarov, S. Treil and A. Volberg, *Bellman function in stochastic control and harmonic analysis.* Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000), 393–423, Oper. Theory Adv. Appl., **129**, Birkhauser, Basel, 2001.

[18] F. NAZAROV, S. TREIL, AND A. VOLBERG, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. of Amer. Math. Soc., **12**, (1999), no. 4, 909-928.

[19] S. PETERMICHL, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris, Sér. I Math., **330**, (2000), no. 6, pp. 455-460.

[20] S. PETERMICHL, A. VOLBERG, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J., **112** (2002), no. 2, pp. 281-305.

[21] S. PETERMICHL, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic.* Amer. J. Math. 129 (2007), no. 5, 1355–1375.

[22] S. PETERMICHL, *The sharp weighted bound for the Riesz transforms.* Proc. Amer. Math. Soc. 136 (2008), no. 4, 1237–1249.

[23] S. PETERMICHL, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol.* C. R. Acad. Sci. Paris S?r. I Math. 330 (2000), no. 6, 455–460.

[24] C. PÉREZ, S. TREIL, A. VOLBERG, *On A_2 conjecture and corona decomposition of weights*, arxiv1005.2630.

[25] C. PÉREZ, S. TREIL, A. VOLBERG, *A_2 conjecture: reduction to a bilinear embedding, splines and less smooth Calderón–Zygmund kernels*, Preprint, pp. 1–17, July 3, 2010.

[26] A. REZNIKOV, S. TREIL, A. VOLBERG, *A sharp estimate of weighted dyadic shifts of complexity* 0 *and* 1, Preprint, September 2010, pp. 1–9.

[27] A. REZNIKOV, A. VOLBERG, *Random "dyadic" lattice in geometrically doubling metric space and A_2 conjecture*, Preprint arXiv:1103.5246, pp. 1–10.

[28] V. VASYUNIN, A. VOLBERG, *Two weight inequality: the case study*, St. Petersburg Math. J., 2005?

[29] A. VOLBERG, *Calderón–Zygmund capacities and operators on non-homogeneous spaces*, CBMS conferences, AMS, vol. 100, 2003, 165pp.

[30] J. WITTWER, *A sharp estimate on the norm of the martingale transform.* Math. Res. Lett. 7 (2000), no. 1, 1–12.

DEPT. OF MATHEMATICS, UNIV. OF WISCONSIN, MADISON

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824, USA